Quarterly Progress Report

Duration: January 1995 — March 1995
ARPA Order Number: A423
Contract/Grant Number: N00014-93-1200
Agent: Office of Naval Research
Contract Title: Outdoor Landmark Recognition Using Hybrid Fractal Vision System and Neural Networks

Organization: North Carolina State University
Box 7911
Dept. of Electrical and Computer Engineering
Raleigh, NC 27695-7911
Principal Investigator: Prof. Ren C. Luo
Phone: (919)515-5199
FAX: (919)-515-5523
e-mail: luo@eos.ncsu.edu

Executive Summary

A hybrid fractal vision system is being developed for landmark detection and recognition in natural scenes. At the current quarter of research, a reconfigurable neural network is being designed to recognize landmarks. The fractal model detected the landmarks for cluttered images, and the neural network would recognize those landmarks. A brief description of the theoretical design of this Reconfigurable Neural Network is given here. Also, some of the initial results obtained by testing the neural network on real image data are included with this report. A new learning method is also being developed and briefly reported here.

Automatic recognition systems can be useful in both military and commercial domains. Tasks such as military surveillance, automatic target recognition, automatic vehicle navigation, material handling, inspection, data compression/decompression, autonomous robot navigation, etc are some of the practical issues directly enhanced by automatic and robust vision systems.

In the next report period, this reconfigurable neural network design would be further developed, the designed network would be thoroughly tested, and the performance of the system would be evaluated. In addition, other image processing techniques would also be evaluated in both spatial and frequency domains.

1 The Reconfigurable Neural Network

The Reconfigurable Neural Network designed for object and landmark recognition is a self organization model for its ability to adopt to new input patterns without the need for retraining. In this model, no feature extraction from the landmarks is performed. Thus less computational demand simplify the algorithm. By representing the data in continuous pattern, most of the information in the traffic signs can be retained. This is a perspective robustness of the approach, a fact which is evident in our initial experiments, but pending further experimental validation.

The weight space of the neural network is flooded with some additional neurons to begin the training process. The additional neurons in the weight space provides two major attributes. First of all, multiple neurons can be used for classes having a high variance of models, and secondly, the free neurons can be assigned to new classes. Critical issues in neural network domain such as the learning rate and the performance are observed carefully during our research and design process.

2 Learning Rules for the Reconfigurable Neural Network

The basic and the most important features for learning and updating the Reconfigurable Neural Network designed are presented briefly in this section. This network is a self organizable entity, a concept which can be understood by visualizing each neuron as being at the center of the N-dimensional hyperspace with all the examples in a particular class available to the neuron. At the beginning of the training, these neurons are placed randomly in the hyperspace, but they move toward the centers of the classes during the training of the network. If the variance of a certain class is high, then the certain

neurons represent only a portion of that class, while the additional neurons available can be assigned to represent those other members or even their subgroups.

For mathematical explanation purposes, assume the input vector to be $\vec{x}$ and the weight vector to be $\vec{w}$. Also symbolize $\vec{x^{\perp}}$ to be a vector perpendicular to $\vec{x}$. The self organization learning is based on a winning concept, meaning that the maximum value of the dot product of the weight and input vectors, $< \vec{w}.\vec{x} >$ is representative of the neuron most closely aligned with the input. The update rule is expressed by the following:

$$\vec{w_k}(n) = \vec{w_k}(n-1) - \alpha < \vec{w_k}.\vec{x^{\perp}} > \vec{x^{\perp}}$$

In the above equation, α is the learning rate. An effective way to find $\vec{x^{\perp}}$ is to initialize $\vec{x^{\perp}}$ to random values, and then take the dot product $< \vec{x}.\vec{x^{\perp}} >$ over all of them. If the two vectors are normal to each other, the dot product should be zero. the mathematical expression in the summation becomes,

$$< \vec{x}.\vec{x^{\perp}} > \quad >= \sum_{i=1}^{N-1} x_i x_i^{\perp} + x_N x_N^{\perp}$$

For normal vectors, as mentioned, the left hand side is zero, and therefore, the normal vector to the input can be calculated to be

$$x_N^{\perp} = -\frac{\sum_{i=1}^{N-1} x_i x_i^{\perp}}{x_N}$$

This is also known as the projection learning method for neural networks.

The learning rate of the neural network has been looked at carefully to set the mathematical basis of the rate at which the neurons update themselves. We need to examine whether the updated neuron would ever yield to an optimal solution. The optimal solution for the neural network is considered to be an alignment of the neuron with the input. During the training process, the learning rate would be gradually reduced. Let

the learning rate at the i^{th} epoch be α_i. Also, let γ_0 represent $< w_0.x^{\perp} >$, a quantity which represent the projection of the initial neuron along the normal of the input. Assuming those symbols, a set of updating weights of the neurons is generated. The update sequence falls into a certain pattern. This pattern can be observed in the following equations:

$$
\begin{aligned}
w_0 &= w_0 \\
w_1 &= w_0 - \alpha_1 < w_0.x^{\perp} > x^{\perp} \\
w_2 &= w_1 - \alpha_2 < w_1.x^{\perp} > x^{\perp} \\
&= w_0 - \alpha_1 < w_0.x^{\perp} > x^{\perp} \\
&\quad -\alpha_2 < (w_0 - \alpha_1 < w_0.x^{\perp} > x^{\perp})x^{\perp} > x^{\perp} \\
&= w_0 - [\alpha\gamma_0 + \alpha_2\gamma_0 - \alpha_2\alpha_1\gamma_0]x^{\perp} \\
w_3 &= w_2 - \alpha_3 < w_2.x^{\perp} > x^{\perp} \\
&= w_0 - [\alpha_1 + \alpha_2 + \alpha_3 - \alpha_2\alpha_1\alpha_3\alpha_1 - \alpha_3\alpha_2 + \alpha_3\alpha_2\alpha_1]\gamma_0 x^{\perp}
\end{aligned}
$$

Thus, the n^{th} updating rule can be expressed as,

$$
w_n = w_0 - \gamma_0[\sum_{i=1}^{N} \alpha_i + \Pi_{i=1}^{N}\alpha_i - \sum_{i=1}^{N} \sum_{j=1,j\neq i}^{N} \alpha_i\alpha_j]x^{\perp}
$$

The learning rate, α is a monotonically decreasing function.

3 Initial Experimental Results

The underdeveloped reconfigurable Neural Network was tested on some typical traffic signs. A list of the four segmented traffic signs are shown in Figure 1. The network learned and recognized signs with sufficient accuracy. A qualitative comparison between this projection learning and the popular Kohonen learning is shown in Figures 2 and 3.

This graph shows that the projection learning has less average error compared to that of the Kohonen learning method as more iterations are made. Also, the input neuron similarity increases more for the projection learning, compared to that of the Kohonen learning, with greater number of iterations.

4 Current and Future Activities

An extensive testing of the Reconfigurable Neural Network using these landmarks are currently being performed. Comparative studies are being conducted for this method to other popular learning methods. The issues of reconfigurability, learning rate, and performance are also included in the research activities. Once these studies are completed, they would be presented in graphical and tabular fashions in the future reports. Last but not least, in addition to the development of this hybrid vision system, explorations of other image processing techniques, both in the spatial domain and the frequency domain, have been initiated recently for object modeling and recognition. Some of the new developments could indeed help design more robust modeling schemes and recognition algorithms. All of the progress would be reported as they are developed in the future.

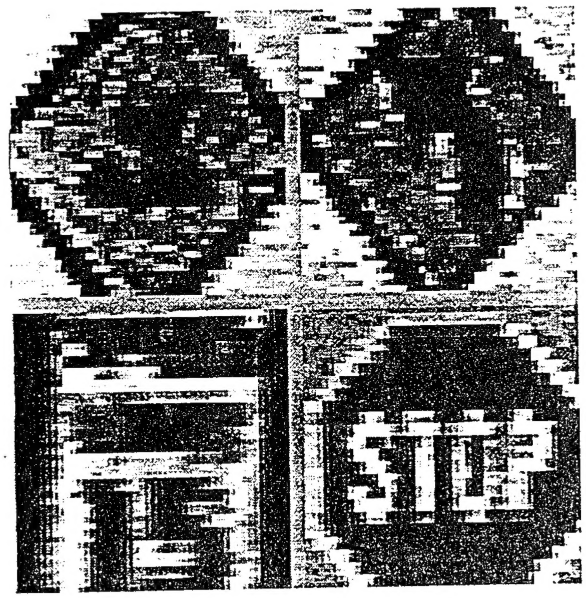

Figure 1: *Examples of traffic signs. Clockwise from top-left: Do Not Enter, Left Turn, Stop, and Speed Limit 15.*

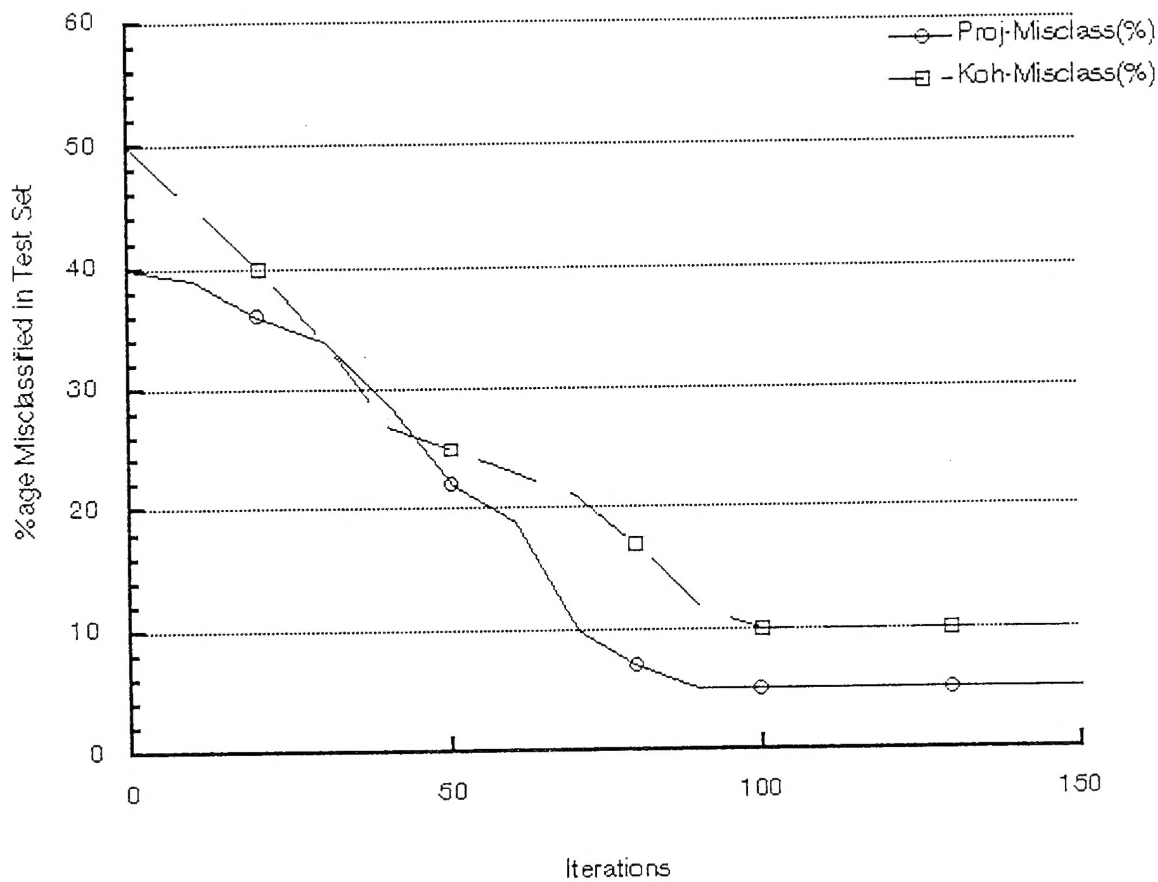

Figure 2: *Comparison of test set misclassification errors: kohonen learning versus projection learning.*

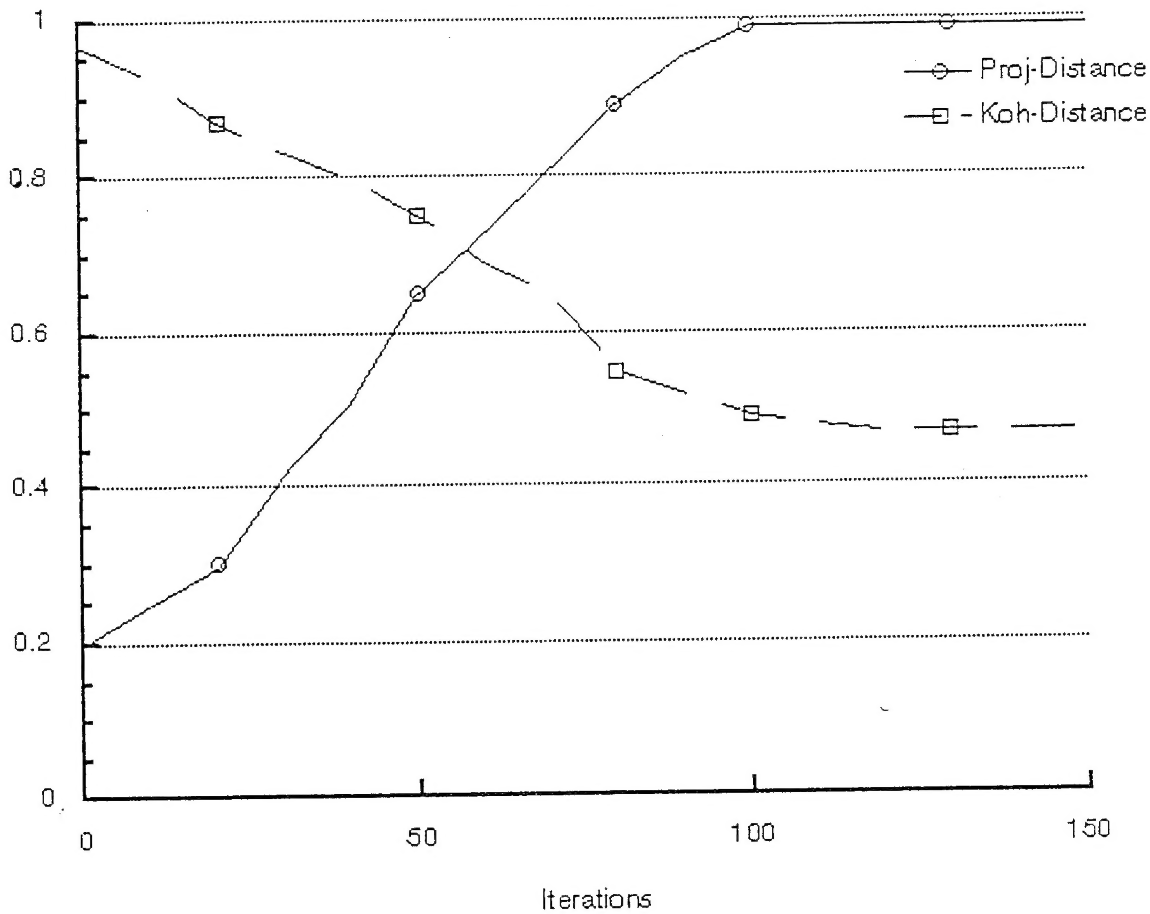

Figure 3: *Comparison of input neuron similarity: kohonen learning versus projection learning.*